DOCUMENT RESUME

ED 120 675 CS 002 505

AUTHOR Powell, William R.
TITLE Monitoring Reading Behavior: Criteria for Performance.
PUB DATE 76
NOTE 22p.; Paper presented at the Annual Meeting of the International Reading Association (21st, Anaheim, California, May 1976)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage
DESCRIPTORS Elementary Education; *Evaluation Criteria; *Informal Reading Inventory; Measurement Instruments; Reading Ability; Reading Achievement; *Reading Diagnosis; *Reading Instruction; Reading Research; Reading Skills; *Word Recognition

ABSTRACT

Effective use of the informal reading inventory (IRI) depends upon the criteria used in determining the functional reading levels and more specifically the word recognition criteria employed in describing acceptable limits of oral reading behavior. The author of this paper looks at the diverse sets of criteria commonly used, the problems associated with these standard approaches, and the two different sets of criteria for word recognition error ratios to be offered for each condition under which the perception data was obtained. A rationale for each of these sets of criteria is presented for each assessment condition as they are developed within a partial theoretical framework. Emphasis is placed on clarifying the criteria problems connected with the informal reading inventory and the teaching and clinical practice which are effected by the evidence offered. (Author/RB)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

William R. Powell
University of Florida
College of Education
186 Norman Hall
Gainesville, Florida 32611

ED120675

MONITORING READING BEHAVIOR: CRITERIA FOR PERFORMANCE

Presented at the

International Reading Association Convention

Anaheim, California

at the Session

"Diagnostic-Prescriptive Reading Instruction Strategies"

Wednesday May 12, 1976

2:00 - 4:45 p.m.

Disneyland Hotel - N. Ballroom #2

CS 002 505

MONITORING READING BEHAVIOR: CRITERIA FOR PERFORMANCE

William R. Powell
University of Florida

Since the beginning of teaching - whether student - or subject matter-oriented or possessing some combination of both, the automatic teacher (Powell, 1972) has endeavored to reach that perfect point in teaching where what the student encounters as new learning in the classroom or clinic is like Baby Bear's porridge -- neither too hot nor too cold. Sometimes consciously, sometimes subconsciously, (and maybe sometimes unconsciously), the teacher has tried to structure his teaching so that the material to be encountered is not too easy - thus boring, - and not too hard - thus discouraging -- but just challenging enough to be exciting and at a level of difficulty that will enable the student, with a reasonable application of energy, to cope with it and succeed.

In essence, this is the stated problem and any teacher worth his chalk will tell you, this is not easy to do. And unless there is some agreement on what constitutes that appropriate level of difficulty and how it is to be determined, then the problem is compounded beyond reasonable understanding and utilization.

The informal reading inventory can assist in solving some of this dilemma or it can further complicate the situation. At its best, it is an effective tool for use in evaluating reading performance and progress. It can be an instrument for use in classroom placement, diagnostic assessment, clinical analysis, and for instant and constant processing of oral reading behavior.

Effective use of the informal reading inventory (IRI) depends upon many factors; mainly, administrative procedures, the definition of error, and the criteria imposed for determining the three functional reading levels - independent, instructional, and frustration.

Purpose

It is the latter factor to which this paper is addressed, namely, the criteria used in determining the functional reading levels and more specifically the word recognition criteria employed in describing acceptable limits of oral reading behavior. What will be presented is a look at the diverse sets of criteria commonly used, the problems associated with these standard approaches, and the, not one, but two different sets of criteria for word recognition error ratios will be offered for each condition under which the perception data was obtained. A rationale for each of these sets of criteria will be presented for each assessment condition as they are developed within a partial theoretical framework. Emphasis will be upon

clarifying the presently confused criteria picture connected with the informal reading inventory and the teaching and clinical practice which are effected by the evidence offered.

State of Affairs

In reality, there are only five different sets of criteria which have been employed in determining the functional reading levels of students: the criteria proposed by Emmett A. Betts (1946), J. Louis Cooper (1952), Nila B. Smith (1959), William R. Powell (1970, 1971a, 1971b), and the criteria implied by and embedded in the various diagnostic oral reading tests, such as the Durrell Analysis of Reading Difficulties, the Gray Oral Reading Tests, the Diagnostic Reading Scales, and the Gates-McKillop Reading Diagnostic Tests. All other criteria in common usage are adaptions or variations of these defined stipulated limits. An examination and comment on each of these five sets of criteria will now follow.

Betts criteria. The first, the best known, and undoubtedly the most widely used word recognition criteria for determining reading levels is that proposed by Emmett A. Betts (1946). Without a doubt, Betts has to be recognized as the "Father of the IRI." While he may have drawn his ideas and hypotheses about what the criteria should be from other sources (Durrell, 1937), it was he who first put it all together in a framework for operation and testing.

Simply stated, Betts set the lower limits for the independent, instructional, and frustration reading levels at 99 percent, 95 percent, and 90 percent, respectively. Translated into word recognition error ratios, the language of classroom teacher, these would be described as allowing for one error for every 100 running words at the independent level, tolerating one error for every 20 consecutive words at the instructional reading level, and declaring one error in every ten running words to be at the frustration level.

Table 1 Betts Criteria

	PP+	Percent
IND.	1/100	99%
INSTR.	1/20	95%
FRUST.	1/10	90%

Betts made no allowance for a differential effect due to age, grade, or difficulty factor of the material. The criteria is and was to be applied unilaterally across all effects. Also, unexplained in the Betts criteria was the wide gap between the lower limits of the instructional level, i. e., 1/20 (to be read- one error for every 20 running words), and the beginning limits of the frustration level, i. e., 1/10.

What about those students whose reading produced scores between 1/11 and 1/19, and how were such findings to be interpreted? However, the most unforgotten factor, and perhaps one of the most significant ones, effecting proper use and interpretation of the Betts criteria was that it was derived by reading silently first and then the assessment was made from the oral reading performance which followed. A practice effect was present. The criteria was determined from a second reading, which was oral, and then generalized to apply to all reading performances, whether first, second, or on subsequent readings. Further, the data which was used to substantiate this set of criteria was only for the instructional reading level of 39 fourth grade students (Killgallon, 1942). The independent level and the frustration level was an assumed extrapolation. In fact, research data clearly aimed to empirically determine those two levels has never been accomplished.

For the past ten years, I was convinced that the Betts criteria was incorrect -- totally and completely. However, as will become evident later in this paper, that was a wrong assumption. Betts was probably correct -- for the population he used and the test conditions by which his instructional reading level was determined. What has been misleading all these years is that reading personnel have applied, generalized, and extrapolated beyond the limits on which the data was based.

Cooper criteria. In his doctoral dissertation, Cooper (1952) investigated the relationship between the percent of word-recognition miscues and the achievement gain in reading. In so doing, Cooper discovered that two sets of criteria were necessary for accurate determination of levels -- one for the primary grades (1-3) and another for intermediate grouped students (4-6). For the primary pupils, the lower limit of the instructional level was 98 percent and for intermediate youngsters, the lower limit was set at 96 percent word pronunciation accuracy. The study contained methodological weaknesses which causes the data and criteria to be held suspect, but Cooper should be commended for his attempt to relate quantitative errors with achievement in reading, and for his realization that a single set of criteria was not appropriate across all age-grade levels.

Table 2 Cooper Criteria

	PP-3		4-6	
IND.	1/100	(99)	1/100	(99)
INSTR.	1/50	(98)	1/25	(96)
FRUST.	1/17	(94)	1/10	(90)

N. B. Smith Criteria. Smith (1959) suggests that the instructional level in grades one, two, and three is best characterized by 80 percent 85 percent accuracy in pronunciation. She makes no mention for criteria above these levels nor if the same criteria are to apply. Further, no evidence was given as to how these criteria were established. One could suspect that good clinical observation and professional intuition led her to her established range. In any instance, the Smith criteria for one readability cluster is remarkably close to the criteria Powell found through the examination of protocol data.

Table 3 N. B. Smith Criteria

	PP-3		4-6
IND.	1/100	(99%)	
INSTR.	1/5	(80%)	?
FRUST.	1/4	(79%)	

Powell criteria. Powell (1970, 1971) began his investigations into the criteria for the designation of the instructional reading level because of the gaps in the Betts' criteria. He began by re-examining Killgallon's data. He adopted the rationale that if a child's comprehension score remained above

70 percent, it could be assumed that the youngster was able to tolerate whatever word recognition patterns accompanied that performance. Accordingly, he turned to the IRI scores in the appendix to Killgallon's study and recomputed the mean and the range of word recognition scores, not against Killgallon's subjective designation of instructional level, but against a comprehension score of 70 percent or better. In contrast to Killgallon's mean of 94.9 percent recognition score or one error in twenty running words, Powell found a mean of 92.8 percent word recognition or one error in fourteen running words.

He continued this line of investigation with 178 full IRI protocols. He examined the distribution of the word recognition scores that accompanied comprehension of 70 percent or better. He found that pupils in grades one and two could tolerate on the average a word recognition score as low as 83 to 87 percent, that is, one error in every seven running words, and still maintain at least 70 percent comprehension. In grades three through five, 90 to 92 percent accuracy in pronunciation accompanied comprehension score of 70 percent or better. In material of grade six or higher, the word recognition percent was found to be about 94 percent. Further, an examination of sixty clinic cases, twelve at each grade level two through six, tended to support the above findings.

Table 4 Powell Criteria: Oral Reading At Sight

	PP-2	3-5	6+
IND.	1/17	1/27	1/36
INSTR.	1/8 - 1/16	1/13 - 1/26	1/18 - 1/35
FRUST.	1/7	1/12	1/17

Three features need to be remembered in interpreting the Powell data. First, the figures cited represent averages, as do the data from other studies. Therefore, one-half of the students read below the stipulated limit and one-half read above it and maintained the required comprehension. Therefore, one could expect that moderation is necessary in interpreting informal data as the lower limits for a given level represent averages, not basal nor arbitrary limits. Secondly, the criteria is applicable to the readability or difficulty level of the material being read and not the age or grade of the youngster. This means that if a sixth grade student is reading from a second grade book, the criteria which would be applied would be the value range for the pre-primer to second grade readability cluster. Thirdly, there is an inverse relationship between the difficulty level of the material and the number of miscues a reader can tolerate. The easier the material, readability-wise, the more error that can be processed and

still maintain an acceptable level of comprehension; the more difficult the material, the fewer the number of miscues that can be tolerated.

Provided the Powell assumption is correct, and there is further evidence by Dunkeld (1971) to suggest that it is, the criterion providing the dividing line between the instructional and the frustration reading level is a differential one depending upon the difficulty of the passage read. Powell's criteria is more in line with the suggestions by Smith, but in conflict with the data from Betts and Cooper. However, the Powell criteria is to be applied to oral reading at sight, while the Betts' criteria really is only applicable to a restricted range when oral reading has first been preceded by a silent reading. The differential function is in agreement with the suggested findings by Cooper.

Diagnostic test criteria. Implicit in the use of many of the more commonly used diagnostic oral reading tests is a set of criteria embedded in the norming process for each respective test. Powell and Dunkeld (1971b) have shown that the criteria implicit in the level determinations of five reputable, established oral reading tests are not at all in accord with either the Betts or the Cooper criteria. The data from the diagnostic tests tend to support the criteria position of Powell and his differential standards at different passage levels.

Table 5 WORD RECOGNITION ERROR RATIOS BY EIGHT SETS OF CRITERIA

Criteria Levels	Powell	Spache	Durrell	Gilmore	Gray	Gates-McKillop	Betts Killgallon	Cooper
P		1/4	1/3	1/3	1/7		1/20	1/50
1^2	1/6	1/5		1/5	1/8		1/20	1/50
2^1		1/8	1/8	1/6	1/11		1/20	1/50
2^2	1/8	1/7	1/9			1/2	1/20	1/50
3^1		1/10		1/8	1/11		1/20	1/50
3^2	1/11	1/13	1/12			1/3	1/20	1/50
4	1/13	1/15	1/13	1/11	1/10	1/4	1/20	1/25
5	1/12	1/16	1/16	1/13	1/11	1/6	1/20	1/25
6	1/17	1/16	1/18	1/14	1/9	1/6	1/20	1/25
7		1/16	1/17	1/18	1/10	1/6	1/20	
8		1/18		1/20	1/9	1/6	1/20	

Reduction theory

The present discussion and evidence does not speak to the issues of the limits of the value ranges for all three functional reading levels, especially the independent reading level; nor does the literature concerning the IRI have any attention given to the theory of processing word recognition errors in the acquisition of a reading repertoire.

In 1973, Powell proposed a "reduction theory" which would explain the processing differences in word recognition scores obtained under differing sets of conditions, i.e., oral at sight and oral reading preceded by silent reading. He surmised that a child's knowledge about the redundancy of print and language with his previous skill acquisition should allow hir. to make logical predictions, thus permitting him to precorrect or reduce his miscues on a second reading.

When an individual reads orally, the reduction of uncertainty is expressed in the reduction of errors -- a quantitative reduction. While it is true that a slight change in word order (a transposition) or a simple substitution may not change the final message significantly, such miscues do not contrubute to the total reduction of the uncertainty of the reading situation. Even slight errors result in modification of the original message and therefore are significant in the total reading process.

This is not to say that reading must be perfect or exact to perceive the message. Such an assumption would be unfounded as the above would indicate. Fortunately, there is enough redundancy in language to allow for errors to occur and the message still to be transmitted. However, given enough errors or miscues within a normal range of values, the probabilities are that the information flow is reduced as the value range is exceeded.

The criteria suggested by Powell is probably suggestive, but not definitive, of the normal error range which can be anticipated for the first (oral) reading of a selection. Quantitatively, the number of errors tolerated is a differential function related to the readabiltiy of the material. Should a reader exceed the normal range of error values, he will have to make more corrections in order not to exceed the redundancy factors of the language and thereby runs the risk of not being able to deal with its unpredictability.

However, if a youngster reads the material a second time as well he might in an instructional setting, then the error range values should change due to the practice effect which gave the pupil an opportunity to achieve precorrection. In a normal developmental teaching of a lesson, the student is (and should be) directed to read the passage first silently for a purpose(s). Out of the discussion that follows, the student might be asked to read selected parts of the material

again -- either silently or orally. If oral reading follows, and the teacher wishes to monitor the oral performance against some guidelines, then these guidelines used should take into account this reduction factor.

For instructional purposes, preliminary research data and clinical observations would suggest that this reduction figure might be 25 to 33 percent. That is, an individual's first reading of a given selection will likely reduce his errors on the second reading of the same material by about one-fourth to one-third. Table 6 reflects this reduction factor for oral after silent reading first and is constructed on the 33 percent precorrection figure. L. L. Smith (1974) has previously constructed a table using a 25 percent reduction factor.

If this assumption is correct, an inspection of the table will reveal that for the second reading the range of error values would be 91.7 percent or one error in every twelve running words, 95 percent or one error in every twenty running words, and 97.3 percent or one error for every twenty-seven consecutive words for readability cluster, respectively.

Table 6 Powell Criteria: Oral After Silent Reading First

	PP-2	3-5	6+
IND.	1/17	1/27	1/36
INSTR.	1/12 - 1/16	1/20 - 1/26	1/27 - 1/35
FRUST.	1/11	1/19	1/26

This criteria table would be the one the classroom teacher would normally use in the conduction of a developmental reading lesson because it accounts for the reduction or precorrection factor that should have occurred during the first reading. If such reduction did not occur, then the teacher might be well served to make a diagnostic analysis of the lesson process to see what happened and why such precorrection did not happen. The student should always be spiraling upward and the reduction factor helps explain that upward movement.

One very interesting sidenote is apparent in comparing Table 4 with Table 6. The lower limit of the three to five readability cluster on the second reading, using a 33 percent reduction factor, is 95 percent or one error in every twenty running words. This is precisely what Betts suggested in the 1940's when he determined his criteria by and through the use

of a second reading. Therefore, it is likely that the Betts criteria for the population (fourth) he used and conditions (silent before oral) was correct and is correct for the designation of instructional reading level within those limitations. It only needed to be tempered with the proper limits of interpretation and not generalized beyond the data which supported it — perhaps a lesson in research to all of us.

Independent reading level

Should a second re-reading (this would be the third time) occur, which is likely in many instructional instances, this third reading will likely reduce the number of perception errors by another 10 to 17 percent. Further readings are not likely to make significant changes in the quantitative factor as reading performance on a given piece of material stabilizes when the reduction values reach those upper limits.

Table 7 REDUCTION FACTOR

FIRST READING	25%	-	33%
SECOND READING	10%	-	17%
TOTAL REDUCTION		50%	

If this observation is true, then the total reduction percentage of a given selection, unless it is overlearned, will be about 40 to 50 percent. If you apply this 50 percent figure to the baseline data suggested by research for the dividing line between the instructional and frustration levels, then you can determine quantitatively the dividing line between the instructional reading level and the independent level. Thus one error in eight running words, for the readability levels preprimer through two, reduced 50 percent equals one error for every 16 running words which would be the upper limit for the instructional level at that readability cluster. Therefore, one error in every seventeen running words would be the beginning point or lower limit of the independent level for material within that difficulty range. The same computation process for each of the other two readability clusters produce the upper limits for each, respectively.

Individualizing instruction

Since the error value ranges represent averages, this would imply that every student has his own unique error value point which is most likely to be within close approximation of the central figure. However, it could vary above or below the reported mean.

If the baseline or lower limit of the line between the instructional and frustration level for a given student were determined, then by applying the one-fourth to one-third reduction factor to his baseline data, his teacher would know what to expect of each child after his first silent reading. His performance could be monitored by the use of the "finger-count technique" devised by Powell. Then the teacher could make an instant diagnosis as to progress of each pupil each time he performs orally. A 50 percent reduction percentage could be computed for each student and the lower limits of his independent level would be established and the basis for guidance in selecting books for recreational reading would be apparent.

Further, it is likely that each individual has his own precorrection or reduction ratio. If this is so, and it is determined and applied to his baseline data at the lower limits of his instructional level, then it would be truly possible to totally and completely individualize.

Then and only then will we, as teachers, serve porridge which is neither too hot nor too cold -- but just right for exciting and challenging new learnings.

REFERENCES

Betts, E. A. Foundations of Reading Instruction. New York: American Book Co., 1946.

Busboom J. B. An Investigation of the Relationship Among Various Testing Techniques on An Informal Reading Inventory. Unpublished doctoral dissertation, University of Illinois, 1973.

Cooper, J. L. The Effect of Adjustment of Basal Reading Materials on Reading Achievement. Unpublished doctoral dissertation, 1952.

Dunkeld, C. G. The Validity of the Informal Reading Inventory for the Designation of Instructional Reading Levels: A Study of the Relationships Between Children's Gains in Reading Achievement and the Difficulty of Instructional Materials. Unpublished doctoral dissertation, University of Illinois, 1970.

Durrell, D. D. Individual Differences and Their Implications with Respect to Instruction in Reading. In G. M. Whipple (ed.), Teaching of Reading: A Second Report. The 36th Yearbook of the National Society of the Study of Education, Part I. Bloomington, Ill,: Public School Publishing Co., 1937. Pp. 325-365.

Killgallon, P. A. A Study of Relationships Among Certain Pupil Adjustments in Reading Situations. Unpublished doctoral dissertation, The Pennsylvania State University, 1942.

Powell, W. R. The Effective Reading Teacher. The Reading Teacher, 1972, 25, 603-67.

Powell, W. R. Reappraising the Criteria for Interpreting Informal Inventories. In D. Boer (ed.), Reading Diagnosis and Evaluation. Newark: International Reading Association, 1970. Pp. 100-109.

Powell, W. R. Validity of the Instructional Reading Level. In R. E. Leibert (ed.), Diagnostic Viewpoints in Reading. Newark: International Reading Association, 1971a. Pp. 121-133.

Powell, W. R. Acquisition of a Reading Repertoire. Library Trends, 1973, 2, 177-196.

REFERENCES (cont.)

Powell, W. R. & Dunkeld, C. G. Validity of the IRI Reading Levels. Elementary English, 1971b, 2, 637-642.

Smith, L. L. The Use of Oral Language by the Diagnostic Teacher. In D. C. Waterman & V. M. Gibbs (Eds.), Oral Language and Reading. Proceedings of the 1973 Third Annual Reading Conference. Terre Haute: Indiana State University, 1973. Pp. 53-58.

Smith, N. B. Graded Selections for Informal Reading: Diagnosis for Grades 1 through 3. New York: New York University Press, 1959.